

AMENDMENT TO THE CLAIMS

1. (currently amended) A computer readable medium including instructions readable by a computer, which when implemented, cause the computer to classify data comprising the steps of:

receiving labeled data;

receiving unlabeled data;

constructing a first classifier and a second classifier using the labeled data;

performing uncertainty reduction comprising:

selecting instances from the unlabeled data that are uncertain with respect to the first classifier; and

selecting instances from the unlabeled data that are uncertain with respect to the second classifier;

labeling the instances uncertain to the first classifier using the second classifier to form a first labeled set;—and

labeling the instances uncertain to the second classifier using the first classifier to form a second labeled set; and

forming a classification module for labeling data based at least in part on the first classifier and the second classifier.

2. (original) The computer readable medium of claim 1, wherein receiving labeled data includes receiving data assigned to classes and wherein receiving unlabeled data includes receiving data capable of being assigned to classes.

3. (original) The computer readable medium of claim 2, and further comprising reconstructing the first and second classifiers using at least the first labeled set.

4. (Previously Presented) The computer readable medium of claim 1, wherein selecting instances uncertain to the first classifier includes selecting instances of the unlabeled data as a function of uncertainty.

5. (original) The computer readable medium of claim 4, wherein selecting instances as a function of uncertainty includes calculating probabilities.

6. (original) The computer readable medium of claim 5, wherein calculating probabilities includes calculating probabilities that the first classifier is unable to label some instances of the unlabeled data.

7. (Previously Presented) The computer readable medium of claim 4, and further comprising calculating at least one value of uncertainty used to select the instances uncertain to the first classifier.

8. (Previously Presented) The computer readable medium of claim 7, wherein selecting instances uncertain to the first classifier includes selecting instances having uncertainty values relative to a predetermined threshold.

9. (Previously Presented) The computer readable medium of claim 8, wherein selecting instances having uncertainty values includes selecting instances having uncertainty values below the predetermined threshold.

10. (cancelled)

11. (Previously Presented) The computer readable medium of claim 1, and further comprising exchanging information between the first labeled and the second labeled set to form at least one shared set.

12. (original) The computer readable medium of claim 11, and further comprising reconstructing the first and second classifiers using the at least one shared set.

13. (Previously Presented) The computer readable medium of claim 1, wherein labeling the instances uncertain to the first classifier includes assigning instances in the first labeled set to a first set of classes, and wherein labeling the instances uncertain to the second classifier includes assigning instances in the second labeled set to a second set of classes.

14. (original) The computer readable medium of claim 13, wherein the first set of classes and the second set of classes are the same.

15. (original) The computer readable medium of claim 13, wherein the first set of classes and the second set of classes are different.

16. (original) The computer readable medium of claim 13, wherein the first set of classes and the second set of classes are mutually exclusive.

17. (original) The computer readable medium of claim 13, wherein the first set of classes and the second set of classes overlap by having classes in common.

18. (Previously Presented) The computer readable medium of claim 1, and further comprising selecting unlabeled data that is relatively certain with respect to the first classifier.

19. (Previously Presented) The computer readable medium of claim 18, wherein selecting relatively certain unlabeled data includes selecting the most certain unlabeled data with respect to the

first classifier.

20. (Previously Presented) The computer readable medium of claim 18, wherein selecting instances uncertain to the first classifier includes selecting data in the relatively certain data.

21. (Previously Presented) The computer readable medium of claim 18, wherein selecting instances uncertain to the first classifier includes selecting data not in the relatively certain unlabeled data.

22. (currently amended) A computer readable medium including instructions readable by a computer, which when implemented, cause the computer to classify data comprising the steps of:

constructing a first classifier and a second classifier  
using received labeled instances;

using the first classifier to select unlabeled instances  
that are certain with respect to the first classifier;

selecting instances uncertain with respect to the second  
classifier data from among the instances certain with  
respect to the first classifier to form a first set of  
unlabeled instances;

using the second classifier to select unlabeled instances  
that are certain with respect to the second classifier;  
and

selecting instances uncertain with respect to the first  
classifier from among the instances certain with  
respect to the second classifier to form a second set  
of unlabeled instances-; and

forming a classification module for labeling data based at  
least in part on the first classifier and the second  
classifier.

23. (original) The computer readable medium of claim 22, and further comprising:

labeling the first set of unlabeled instances using the first classifier to form a first labeled set;  
labeling the second set of unlabeled instances using the second classifier; and to form a second labeled set;  
adding the first labeled set and the second labeled set to the received labeled instances to form an augmented set.

24. (original) The computer readable medium of claim 23, and further reconstructing the first classifier and the second classifier using the augmented set.

25. (original) The computer readable medium of claim 24, wherein reconstructing the first and the second classifiers includes iteratively reconstructing the first and the second classifiers for each class of unlabeled instances.

26. (original) A data classification system comprising the computer readable medium of claim 1.

27. (currently amended) A method of training a classifier, the method comprising:

receiving labeled data;  
receiving unlabeled data;  
constructing a first classifier and a second classifier using the labeled data;  
using the first classifier to select some of the unlabeled data that is certain with respect to the first classifier;  
selecting unlabeled data uncertain with respect to the second classifier from among the unlabeled data that is certain with respect to the first classifier to form a first set of unlabeled data; and  
using the second classifier to select some of the unlabeled data that is certain with respect to the second

classifier;~~and~~  
selecting unlabeled data uncertain with respect to the first  
classifier from among the unlabeled data that is  
certain with respect to the second classifier to form a  
second set of unlabeled data~~;~~ and  
forming a classification module for labeling data based at  
least in part on the first classifier and the second  
classifier.

28. (original) The method of claim 27 wherein receiving unlabeled data includes receiving data capable of being assigned to classes, and wherein receiving labeled data includes receiving data assigned to classes.

29. (original) The method of claim 28, and further comprising reconstructing the first and the second classifier for each class.

30. (original) The method of claim 27, and further comprising:  
applying the first classifier to the first unlabeled set to  
form a first labeled set; and  
applying the second classifier to the second labeled set to  
form a second labeled set.

31. (original) The method of claim 30, and further comprising augmenting the received labeled data with the first labeled set and the second labeled set to form an augmented set.

32. (original) The method of claim 31, and further comprising using the augmented set of labeled data to retrain the first classifier and the second classifier to form a retrained first classifier and a retrained second classifier.

33. (original) The method of claim 30, wherein applying the first classifier to the first unlabeled set includes calculating

probabilities that the first classifier is unable to assign some unlabeled data to classes.

34. (original) The method of claim 33, wherein calculating probabilities includes calculating values of uncertainty.

35. (original) The method of claim 34, wherein calculating values of uncertainty includes calculating values of uncertainty relative to a predetermined threshold.

36. (currently amended) A method of assigning information into classes, the method comprising:

receiving labeled data;

receiving unlabeled data;

constructing a first classifier and a second classifier with the received labeled data;

selecting a portion of unlabeled data that is uncertain for the first classifier;

assigning classes to the portion of unlabeled data using the second classifier to form a first labeled set;

selecting a portion of unlabeled data that is uncertain for the second classifier;~~and~~

assigning classes to the portion of unlabeled data uncertain for the second classifier using the first classifier to form a second labeled set~~;~~ and

forming a classification module for labeling data based at least in part on the first classifier and the second classifier.

37. (cancelled)